INDICADORES

THE NEED FOR END-USER CUSTOMIZATION OF THE JOURNAL-SETS OF THE SUBJECT CATEGORIES IN THE *SCIMAGO JOURNAL RANKING DATABASE* FOR MORE APPROPRIATE LEAGUE LISTS. A CASE STUDY FOR THE LIBRARY & INFORMATION SCIENCE FIELD

Péter Jacsó

Nota: Este artículo puede leerse traducido al español en:
http://www.elprofesionaldelainformacion.com/contenidos/2013/sept/12_esp.pdf

Péter Jacsó is a professor in the *Department of Computer and Information Sciences* at the *University of Hawaii*. He has been teaching courses on library automation, digital libraries, advanced database searching, and computerizing serials management. He has years of practical experience in running computer centers, managing library and information services, and evaluating database software and content. He has published extensively on these topics in journals, conference proceedings, and books, and received awards for his teaching and database reviews.
http://orcid.org/0000-0003-1409-2404

Univ. of Hawaii, Dept. of Information and Computer Sciences, Library and Information Science Program
2550 The Mall, Hamilton Bldg. # 2-I. Honolulu, HI 96822, USA
http://www.jacso.info
jacso@hawaii.edu

Abstract

The open access *SCImago Journal & Country Rank* (*SJR*) service -built from data in the subscription-based *Elsevier*'s *Scopus* database- offers bibliographic information and bibliometric indicators for nearly 20,000 journals and other serial publications for 1996-2011. Journals are assigned to 300+ pre-defined subject categories and 26 broad subject areas to produce league lists. The set of 134 journals assigned to the Library and Information Sciences (LIS) subject category and a sample set of 50 other LIS journals, scattered across several other categories, were analyzed in the latest edition (*SJR*-2011). Recommendations are made on how *SJR* could be enhanced by its developers to include simple customization options in the interface to facilitate a more efficient look-up and more valid comparison of the standing of journals and other serials and to reflect the prevailing preferences at the institutions where the rank positions in journal league lists are used to inform career advancement and funding decisions. It is argued that extending the current three-year citation window to five years would be more appropriate for all disciplines, and would also make *SJR* comparable to metrics-based journal league lists produced from the open access *Microsoft Academic Search* (*MAS*), *Google Scholar Metrics for Publications* (*GSMP*), and *Eigenfactor.org* (*EF*) data and the subscription-based *Journal Citation Reports* (*JCR*), which pioneered the idea of a metrics-based journal league list.

Keywords

Journal ranking, League lists, *SCImago Journal Rank*, *Scopus*, Journalogy, Metrics-based publication assessment, Bibliometrics, Scientometrics, Informetrics.

Título: Necesidad de que los usuarios finales puedan personalizar los conjuntos de revistas de las categorías temáticas en la base de datos *SCImago Journal Ranking* para obtener rankings más apropiados. Estudio de caso del campo Biblioteconomía y Documentación

Resumen

El servicio en acceso abierto *SCImago Journal & Country Rank* (*SJR*) –elaborado a partir de la base de datos de pago *Scopus*, de *Elsevier*- ofrece información bibliográfica e indicadores bibliométricos de casi 20.000 revistas y otras publicaciones periódicas en el período 1996-2011. Para producir los rankings, a las revistas se les asignan más de 300 categorías temáticas predefinidas, así como 26 áreas temáticas más amplias En este trabajo se ha analizado el conjunto de 134 revistas asignadas a la categoría Biblioteconomía y Ciencia de la Información (LIS) y una muestra de otras 50 revistas –que también tratan LIS, pero dispersas en varias otras categorías-, todo ello en la última edición *SJR*-2011. Se hacen recomendaciones sobre cómo los desarrolladores del *SJR* podrían mejorar el producto si incluyeran opciones de personalización simples en la interfaz para

Article received on 08-07-2013
Final acceptance: 28-07-2013

facilitar una visualización más eficiente y una comparación mejor de la situación de las revistas y otras publicaciones periódicas. Así también se reflejarían las preferencias de las instituciones para las que la posición de las revistas en el ranking la utilizan para informar la promoción profesional y las decisiones de financiación. Se argumenta que la ampliación de la ventana de citación actual de tres a cinco años sería más adecuada para todas las disciplinas, y también haría *SJR* comparable con los rankings en acceso abierto producidos por *Microsoft Academic Search* (*MAS*), *Google Scholar Metrics for Publications* (*GSMP*), *Eigenfactor.org* (*EF*) y de los *Journal Citation Reports* (*JCR*) –éste de aceso mediante suscripción de pago-, que fue pionero en idear un ranking de revistas basado en métricas.

Palabras clave

Ranking de revistas, *SCImago Journal Rank*, *Scopus*, Estudios de revistas, Evaluación de publicaciones basada en métricas, Bibliometría, Cienciometría, Informetría.

Jacsó, Péter (2013). "The need for end-user customization of the journal-sets of the subject categories in the *SCImago Journal Rank* database for more appropriate league lists. A case study for the Library & Information Science field". *El profesional de la información*, septiembre-octubre, v. 22, n. 5, pp. 459-473.

http://dx.doi.org/10.3145/epi.2013.sep.12

Introduction

SCImago Journal & Country Rank (*SJR*) was the first service to offer an open access alternative –from a different source- to the subscription-based *Journal Citation Reports* (*JCR*) of the *Institute for Scientific Information*, *ISI* (now *Thomson-Reuters*). Among the open access journal league lists, *SJR* is the largest, best designed, and richest in indicators. However, *SJR* does not provide all of the *JCR* features, such as the highly informative lists of citing and cited journals with graphs and tables of the number of citations received from and given to other journals and themselves (i.e. self-citations), their impact factors (if covered by *JCR*) and the yearly distribution of citations. While the *JCR* has been available for the Sciences and the Social Sciences, *SJR* extends its coverage to include the Arts & Humanities journals, –where journals still matter-.

SJR also includes a *Country Ranking* module for assessing the scholarly publication productivity and impact by country, based on the affiliations of authors. There is also a standalone application, the *SCImago Institutional Ranking Service*, based on the authors' institutional affiliations. It would have been better to give each component a unique acronym, such as *SJR*, *SCR*, and *SIR* to avoid confusion. Henceforth *SJR* will be used only for the *SCImago Journal Ranking* component itself.

SJR was launched as a free service in 2007, based on data produced by *Elsevier* for its *Scopus* database, which in turn was meant to compete with the *Web of Science* (*WoS*) database. The deficiencies of *SJR* are inherited from *Scopus*, which still has inconsistent and odd assignment of journals to subject categories, as well as significant gaps in the coverage of many journals. This paper proposes a solution for the former problem, and a follow-up paper will discuss the implications of the latter, both from the perspective of journal ranking and league lists.

Adding some relatively simple software features to *SJR*, as outlined in this paper, for customizing the disciplinary league lists of nearly 20,000 journals could provide a very comprehensive and still easy-to-use resource for renewal and cancellation decisions in collection development. It would also help in the assessment of published research related to promotion, tenure and funding decisions where the standing of journals in which applicants have published is of high importance.

There have been several substantial reviews of *SJR* itself ever since its debut and the interest in its use apparently keeps growing, as *SJR* (and *Scopus*) are compared against other services which are used to create journal, institution and/or country league lists (**Bergman**, 2012; **Delgado-López-Cózar** *et al.*, 2013; **Falagas** *et al.*, 2008; **González-Pereira** *et al.*, 2009, 2010; **Jacsó**, 2010; **Oppenheim**, 2008; **Torres-Salinas** *et al.*, 2010). Especially relevant is the paper where the authors recommend improving the subject classification in *SJR* on the basis of reference analysis (**Gómez-Núñez**, 2011).

> "The interest in the *SJR* use apparently keeps growing"

For the background of metrics-based ranking and comparison of journals, and about the digital sources from which they are created, the following papers can provide further information to the readers (**Bar-Ilan**, 2010; **De-Moya-Anegón** *et al.*, 2007; **Deis**; **Goodman**, 2005; **Dess**, 2006; **Garfield**, 1999; 2005; **Jacsó**, 2007, 2008, 2009a; **Moed**, 2010; **Pudovkin**; **Garfield**, 2004, 2009; **Raj**; **Zainab**, 2012; **Wagner**, 2009).

The content profile of *SJR*

The 2011 edition of *SJR* (*SJR*-2011) has bibliographic information and bibliometric indicators for more than 19,700 serial publications, including scholarly and professional journals, bulletins, newsletters, book series and conference proceedings. For the sake of simplicity, serials and journals will be used synonymously in the rest of the paper as a group term for all these document types.

SJR covers 1996-2011, providing yearly detail from 1999 onward. For 1996, 1997 and 1998 only the 3-year cumulative indicators are given, as the purpose of the *SJR* indicator and

other related indicators has been to report the productivity and citedness of the journals in the previous 3-year period, vis-a-vis the classic 2-year window through the traditional *Journal Impact Factor* (*JIF-2*) introduced by **Eugene Garfield** in print and then microform and digital formats as the annual *Journal Citation Reports* (*JCR*). The only exception for the 3-year citation window in *SJR* is the set of citations/document (citation rate) indicators, which are provided for 2, 3 and 4 year windows, although the last one appears only in the data table and not in the league lists of *SJR*.

For *SJR*-2011 there are 8.85 million master records for documents. The distribution of records by source type is not specified but by running a no-holds-barred search in *Scopus* for 2008-2011 to retrieve all the records, it can be estimated that 80.3% of the records are from academic journals, 1.7% from trade publications, 15% from conference proceedings, and 3% from book series.

By document type, the distribution is 66% journal articles, 6% review articles (of the research literature), and 19% conference papers. These make up the so-called citable items (91%), the term introduced originally for the *JCR*. In reality, some of the other document types (editorials, short notes, letters to the editors, errata) are also cited occasionally, but to a minimum extent. As the assignment of document types has not been accurate and consistent in databases, it is worth mentioning because *SJR* and *JCR* use the distinction in the league tables, *MAS* and *GSMP* do not, and *Eigenfactor* uses only the citable document types for measuring their productivity and impact.

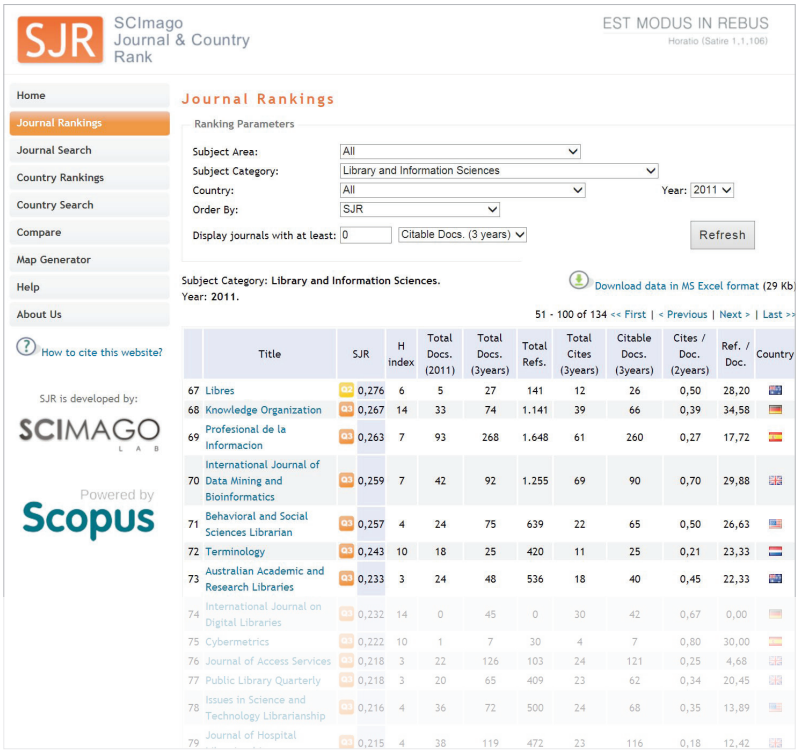


Figure 1a. Excerpt from the league table of journals in the LIS subject category
http://www.scimagojr.com/journalrank.php?category=3309

> *SJR* takes into account a maximum or 33% of the total references which cite the journal itself

The *SJR* league lists can be looked up for each year by simply selecting any year as the "census year" from 1999 to 2011. The citation window is fixed to 3 years. For example, *SJR*-2011 has the census year of 2011 (Y1), and produces most of its indicators by calculating the number of citations received from papers published in 2011 by the documents published in 2008 (Y1-1), 2009 (Y1-2) and 2010 (Y1-3). Some other indicators use a different calculation method. The cited documents indicator refers to those documents which received at least one citation in any year within the assessment time frame of 2008, 2009, 2010 or 2011.

The *SJR* indicator considers not just the number of citations but also the prestige of the citing sources, so the citations received in the year selected by papers published in the previous three years are counted and weighted. This is the same approach used in the *Eigenfactor* database, but the latter automatically excludes all self-citations at the journal level. *SJR* takes into account a maximum of 33% of the total references which cite the journal itself (confirmed by **González-Pereira**, personal communication, 2011). This unusual treatment of journal-level self-citation is not mentioned in the help file, but should be, and prominently, as most users are not aware of this policy.

The *SJR* 2012 edition was expected to be published by the first quarter of 2013. This is not a critical issue from the perspective of the current research, which focused on the most recent edition (*SJR*-2011 as of mid-July, 2013) for subsequent comparison with competing league lists that calculate bibliometric indicators for papers published between 2007-2011, from databases such as the 2011 editions of the *JCR*, *Eigenfactor*, *GSMP*, and *MAS* databases, which offer indicators for –among others- a 5-year window.

Data are imported from the *Scopus* database. If after the import the *Scopus* database is updated with records for documents in its still much needed back-filling process to eliminate or reduce the gaps, these changes are not reflected by the latest *SJR* indicators until its next update. That is the reason that a current search from the *Scopus* database (updated several times a week) may show for 2008-2011 more documents published in a journal, and more citations received by it than reported in *SJR*. The follow-up of this paper will focus on the consistency of the breadth of coverage of LIS journals.

The recommended 5-year citation window

Extending the 3-year citation window to a 5-year period would be a useful enhancement for several reasons. One

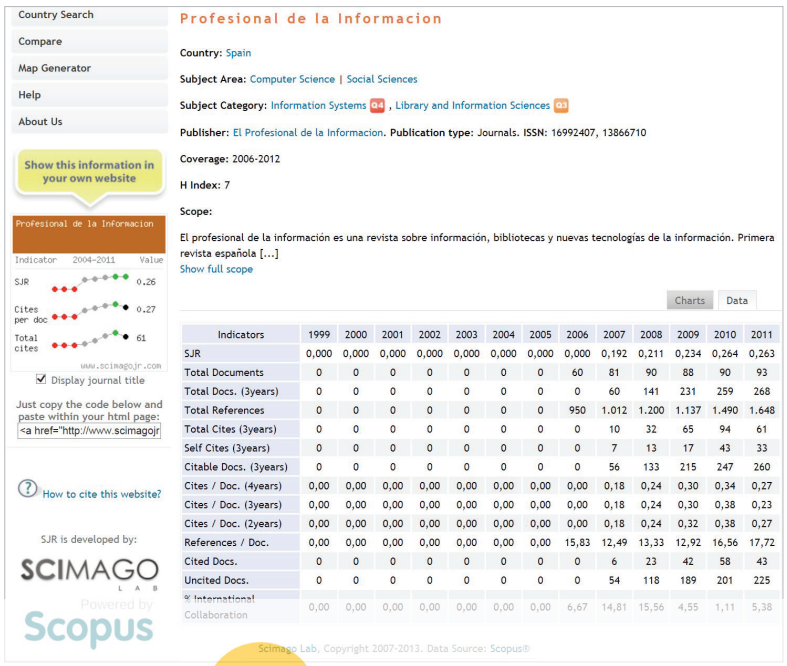


Figure 1b. The data table of a journal with additional indicators about its performance
http://www.scimagojr.com/journalsearch.php?q=16992407&tip=iss

is that the 3-year period is too short in most disciplines. Journals may reach their citation peak by the 3rd year after publication (**González-Pereira *et al.***, 2010) but citations received in the 4th and 5th and subsequent years still could significantly increase most of the indicator scores of the journals' productivity and impact. As we can learn from *JCR*-2011, in the LIS field the aggregated Cited Half-Life was 7.1 years, i.e. the median age of the papers that were cited in the census year. In *JCR*-2012 (which was just released as this manuscript went to press) the Cited Half-Life is 7.3 years). In the Computer Science Information Systems subcategory these indicators were 6.7 and 7.0 for 2011 and 2012, respectively. This could be an argument for using a longer citation window, but the other reasons still make the 5-year window better for comparing databases that offer indicators for 5-year window.

Figure 1c. The impact factors of a journal for a 2, 3 and 4-year citation window
http://www.scimagojr.com/journalsearch.php?q=16992407&tip=iss

The 5-year citation window would be also more appropriate for promotion/tenure application support and post-tenure assessment where the 5-year time span and repeating cycle are very common. All of the above-mentioned databases use a 5-year window (sometimes with other options) versus the original 2-year window used in the classic *JCR*, which I refer to as *JIF-2* for clarity and to distinguish it from the 5-year impact factor (*JIF-5*) added to *JCR* in 2009.

It is recommended to calculate and display in the *SJR* league table and Details table the indicators for a 5-year window for **only** those journals covered by *Scopus* consistently every year between 2007 and 2011 (for *SJR* 2011). This restrictive inclusion principle should also be applied if the developers wouldn't extend the 3-year window. It is to be noted that having comprehensive data for the chosen window (i.e. without gaps at the volume and issue level) can be more important than the choice of the width of the time frame, as will be discussed in the follow-up paper).

The content of the league table for the LIS subject category, a journal Data table and one of the many standalone charts from *SJR*-2011 are illustrated in figure 1a, 1b and 1c. The content, layout and legibility of the league tables is very good. It will be recommended, nevertheless, that in the additional alternate league tables the large logos be moved to another area, in order to accommodate more indicators (currently present in the journal's data table) that are very informational for end-users, such as the ratio of Cited Documents, and the suggested External Citation and Cites/Doc. rate for 5 years.

The competitive context

The journal set of *SJR*-2011 is of almost twice the size of the number of serials covered by the most current edition of *Thomson-Reuters' Journal Citation Reports* (*JCR* 2011): 19,400 versus 8,300 – at first sight. However, the journal set size information must be qualified and interpreted appropriately in the comparable subject areas and categories, rather than just taken at face value. The set of 134 journals assigned to the LIS category in *SJR* was analyzed in detail for coverage along with 50 other, undoubtedly LIS-related, journals scattered across a dozen other subject categories and not included in the LIS subject category.

GSMP limits the league list to the top 20 journals in every subcategory even if it has information about many other journals in the subject category. As usual, *Google* does not reveal the number of journals, let alone the titles included in *GSMP*. The attitude of giving as the fox gives (soup to the stork) – Editor's note: to give something but with

limitations-, has been a trademark of the developers of *Google Scholar*. The two indicators, the 5-year h-index and the median citations, are already calculated automatically for many more journals in order to generate the top-20 list, so displaying the top 50 or even the top 100 journals in LIS would have not required additional effort. Of course, this may give away too much – by *Google*'s standard- about how well or how badly the *GSMP* software crawls and parses digital journal collections.

The *GSMP* service was launched April 1, 2012 (ominously, on April Fools' Day), from a cleaned-up version of *Google Scholar*. To the credit of the *GSMP* developers, the clean-up was good compared to regular *Google Scholar* data, although still not sufficiently reliable, showing duplicate entries with minor differences in accented and other special characters (**Jacsó**, 2012, 2013; **Delgado-López-Cózar** *et al.*, 2012). In the original announcement of the service it was promised that the *GSMP* data would not change for one year ("*Scholar Metrics* are based on our index as it was on April 1st, 2012. For ease of comparison, they are NOT updated as the *Scholar* index is updated."). Actually, *GSMP* was re-launched seven months later after some additional clean-up. The 5-year h-index and the h-5 median indicators are not informative enough without knowing the total number of documents retrieved from and citations matched for the journal by *GSMP*. Again, it would reveal too much about the (in)comprehensiveness of the source coverage in *GSMP*. *Google* developers should have a simple utility to check duplicate entries in league lists, a sort option by journal name, an increased limit to at least 50 journals, and should include the number of documents and number of citations used for each journal to calculate the two indicators.

> "In *SJR* there are 50 undoubtedly LIS-related journals scattered across a dozen other subject categories not included in the LIS subject category"

The *Eigenfactor Service* (*EFS*) –based on *JCR* data- provides information on only 45 of the 83 LIS journals included in the *JCR*-2011 Information Science and Library Science (ISLS) category.
http://www.eigenfactor.org/rankings.php?bsearch=NU&searchby=category&orderby=Eigenfactor

It "adds insult to injury" that some of the most widely known and most influential ISLS journals are excluded. This massively distorts the real rank positions of many of the 45 journals included in the ISLS category by the *Eigenfactor* database. Some of the serials missing from this pseudo-*JCR*/ISLS league list are *ASLIB proceedings*, *Information research*, *Information technology and libraries*, *JASIS&T*, *Journal of information science*, *Journal of information technology*, and *Scientometrics.* The ontology created by the *Eigenfactor* developers is somewhat better, having indicators for 53 journals in its Information Science category (it does not have an **LIS** category), but still much too limited for a sufficiently comprehensive league list of LIS journals: it has only 3 indicators at a precision of 7 decimal places, which are not readily comprehensible by mere mortal end-users. The idea of weighting the citation counts by the prestige of the citing source was a very good one, and the use of percentile bar-charts is very good, but the implementation needs much more spot-checking to minimize non-sensical information.

Figure 2. Journals erroneously assigned to the Infectious Diseases category of the *Eigenfactor* "ontology" in the 2010 edition of the database are still not corrected in that edition as of mid-2013.
http://www.eigenfactor.org/rankings.php?search=118&year=2010&searchby=efcat&orderby=Eigenfactor

The other absurdity in the handling of league lists in the *Eigenfactor* database is the assignment of journals to completely unrelated categories. For example, some of the journals that ended up in the Infectious Diseases subject category within the *Eigenfactor* ontology in the 2010 edition of the database were the following titles: *Austrian journal of earth sciences*, *Education sciences and technology*, *Energy information development*, *Journal of ovonic research*, *Financial research letters*, *Miskolc mathematical notes*, *Platinum metals review*, and *Turkish online journal of education technology*. These were corrected in the 2011 edition after the critical review (**Jacsó**, 2012b), but in mid-2013 they still appear in the 2010 edition of the *Eigenfactor* database.

MAS has a Library Science category with 72 journals in 2013 (62 for the 2007-2011 window). This is a realistic set because the category is LS, not LIS. *MAS* uses the term "current 5-year", so as time passes by for many journals the 5-year time span is shifting. In *MAS* there are some journals already where the 5-year span is from 2008 to 2012. It would have been better from Microsoft to have separate edition(s) for each year, e.g. *MAS*-2010, *MAS*-2011 to avoid ambiguity by the partially shifting 5-year window.

The subscription-based *Scopus* database itself did not create league lists, but started to directly provide journal-level bibliometric data through its Journal Analyzer module, including now the *SJR* and *SNIP* (*Source-normalized impact per paper*) indicators. The latter makes it possible to compare the prestige/impact of journals in different disciplinary areas by correcting for their very different citation patterns. This is highly relevant for some subject areas, such as LIS, where information science journals usually dominate the

1700		COMPUTER SCIENCE Subject area
code	**# of JNs**	**Subject categories**
1701	165	Computer Science (miscellaneous)
1702	114	Artificial Intelligence
1703	127	Computational Theory and Mathematics
1704	78	Computer Graphics and Computer-Aided Design
1705	140	Computer Networks and Communications
1706	204	Computer Science Applications
1707	40	Computer Vision and Pattern Recognition
1708	124	Hardware and Architecture
1709	40	Human-Computer Interaction
1710	132	Information Systems
1711	45	Signal Processing
1712	184	Software

Figure 3a. The suggested hierarchically browsable subject category index for the Computer Science broad subject area

code	# of JNs	Subject categories
3501	93	Dentistry (miscellaneous)
3502	1	Dental Assisting
3503	1	Dental Hygiene
3504	9	Oral Surgery
3505	8	Orthodontics
3506	1	Periodontics
3601	41	Health Professions (miscellaneous)
3602	5	Chiropractics
3603	3	Complementary and Manual Therapy
3604	1	Emergency Medical Services
3605	12	Health Information Management
3606	1	Medical Assisting and Transcription
3607	12	Medical Laboratory Technology
3608	0	Medical Terminology
3609	2	Occupational Therapy
3610	3	Optometry
3611	11	Pharmacy
3612	70	Physical Therapy, Sports Therapy and Rehabilitation
3613	1	Podiatry

Figure 3b. The abuse of the miscellaneous category in Dentistry and the irrational ones in Dentistry and Health Professions with 0-4 journals

code	# of JNs	Subject categories
2001	169	Economics, Econometrics and Finance (miscellaneous)
2002	388	Economics and Econometrics
2003	125	Finance

Figure 3c. *Elsevier*'s miserly and reticent treatment of the Economics, Econometrics and Finance broad subject area by parroting its components as subject categories – except for Econometrics

top positions and library science journals often compete with journals in communication and education in the various databases used for creating league lists.

Surprisingly, *Scopus* does not even use the subject category terms developed by *Elsevier* for the journal title list of –well, *Scopus*. One feature that was not well implemented in the otherwise state-of-the-art *Scopus* software was the summarization and visualization of the journals' performance indicators (**Jacsó**, 2007, 2008, 2009c). Even today, the *Scopus* Journal Analyzer module looks as if it had been an afterthought. This functionality is further limited by the restriction that a maximum of 10 journals can be handled by the Journal Analyzer module.

Subject areas and categories in *SJR*

SJR relies not only on the data content of *Scopus* but also on the classification of the source publications into 26 broad subject areas and 300+ subject categories (which are like subject descriptors of controlled vocabularies in many databases).

Subject areas

Elsevier assigns serials to one or more of the 26 (technically, 27) broad subject areas, ranging from Agricultural and Biological Sciences to Veterinary Sciences. The 27th broad subject area is labeled "Undefined," with 778 items in *Scopus*. Volunteering this information is appreciated, but it is not obvious why these items could not be assigned to one of the 26 other *Scopus* subject areas. *SJR* does not have this category. A review of some of the records in the Undefined category suggests that these entries are simply incomplete records, missing the metadata element that assigns a record to a subject area.

By far the largest number of journals is assigned to the broad subject area of Medicine. There are 5,312 serials assigned to that subject area, about 27% of the actively covered serials for 2008-2011 in *SJR*. On average, journals are assigned to 1.3 broad subject areas. This high share is also reflected by the 2.5 million records within the Medicine subject area, a similar percentage of the 8.8 million total records for documents processed by *Elsevier* for the same time period in *Scopus*, which served as the document base for *SJR*-2011.

Ideally, as the expectation with documents would be evenly distributed among the 26 subject areas. However, the imbalance in the coverage by broad subject areas is quite clear, which in turn suggests some changes. Nursing, Health Professions, Dentistry and Decision Sciences are also listed at the top level among the 26 broad subject areas, in the company of Computer Science, Physics/ Astronomy, Biochemistry/Genetics, and Materials Science. The first three, which -combined- barely exceed 3.6% of the 8.8 million records, could have been assigned to a lower-level subject category, such as Complementary Medicine or Allied Health. Decision Sciences (with fewer than 80,000 documents) could have been one of the lower-level subject categories of the Business, Management and Accounting or the Computer Science subject areas.

Apparently, the design of the subject hierarchy was driven more by the pre-eminent databases of *Elsevier* in medicine and healthcare, such as *EMbase* and *EMcare*, and the *EMtree Thesaurus*, which was enhanced by 1,500 terms for the nursing field alone.

The mean number of journals per subject areas is 954, the median is 606, and there are areas where the number of journals is excessively large, such as Engineering (1,767), Agricultural and Biological Sciences (1,622), Biochemistry, Genetics, and Molecular Biology (1,506).

The number of journals assigned to a subject area does not imply breadth of coverage at the record level. For example, the 3rd largest subject area by number of journals is Arts & Humanities (2,002), but there are only 240,000 records in that broad subject area, which seemed to have been an afterthought to make *Scopus* appealing to Arts & Humanities libraries (**Jacsó**, 2007). On the other hand, there is no *JCR* category for Arts & Humanities, so that is a definite asset of *SJR*, despite its relatively small volume of documents.

Subject categories

Elsevier also identifies, in its publicly accessible journal list, the subject categories assigned to each journal. There are more than 300 subject category terms. A journal may be assigned to several subject categories. No perfect assignment can be expected for nearly 20,000 journals in 26 broad subject areas, but browsing through several hundred journals and checking their assignment to subject categories indicated that the ontology is mostly good, logical and uses common terminology, not the strained and indirect language of many controlled vocabularies. However, there are serious shortcomings for some subject areas and subject categories that must be corrected by *Elsevier* and then updated in *SJR*.

The absurdities in some of the subject category levels –discussed below- may be the reason why *Scopus* does not use these subject descriptors in the records, let alone make them search criteria. Only the 26 very broad areas appear in the *Scopus* records and are offered for searching and filtering the results. These are not appropriate for creating league lists, which require much more specific subject categories.

This is where *SJR* and another similarly appealing and standalone service, the *CWTS Journal Indicators* enter the scene. *http://www.journalindicators.com*

They demonstrate the importance and power of the natural intelligence and bibliometric experience of the members of the *SCImago Research Group* and *Leiden University*'s *Centre for Science and Technology Studies* (*CWTS*). The latter implemented, integrated and at the end of 2012 updated **Henk F. Moed**'s original *SNIP* indicator algorithm (reflecting the significant differences among the citation practices in the various disciplinary fields) to produce an open access league list of nearly 20,000 journals from the *Scopus* database (**Moed**, 2010; **De-Moya-Anegón** *et al.*, 2007; **Waltman** *et al.*, 2012).

One of the most apparent advantages of *SJR* and the *CWTS* services (beyond being open access) is that their designers were aware of the existing classification of the journals covered by *Scopus* into 300+ subject categories and built them into their services, which was not done in *Scopus* itself. These subject area terms make it possible to produce league lists by practical and mostly well-identified research fields, but the vocabulary requires some clean-up.

Assignment of subject categories to journals

There are weaknesses in the subject term assignments in *SJR*, which are inherited from *Elsevier*'s journal list in *Scopus*. The following examples illustrate the major deficiencies in the assignment of journals to subject categories.

Scientometrics, the first journal dedicated to this subject is – correctly- assigned to LIS, along with two other categories: Computer Science Applications, and Computational Theory and Mathematics. *Journal of informetrics* is assigned to 5 subject categories (Applied Mathematics, Computer Science Applications, Management Science and Operation Research, Modeling and Simulation, Statistics and Probability), but not to LIS. Similarly, *Law library journal* is assigned only to the Law category, and *Legal reference services quarterly* only to the LIS category. It would be very appropriate to assign these two journals to both the LIS and Law subject categories. *Journal of digital information* is assigned to the Information Systems category only, and not to LIS, while *Journal of digital information managemen*t is assigned to LIS and two other subject categories.

It is not clear why *College & research libraries* (entered with "and" as the conjunction rather than with the ampersand symbol used by this and hundreds of other journals) is assigned correctly to the LIS subject category, but *College & research libraries news* is assigned to the Education subject category and not to LIS.

Some categories are redundant, as evidenced when one category covers almost all of the journals also assigned to the other. For example, the subject category of Gerontology (with 10 journals) overlaps Geriatrics and Gerontology (with 84 journals). Only 3 journals in the former category are not assigned to the latter. In addition, there is a subject category of Aging. This is a symptom of the inconsistent assignment of journals to subject categories by *Elsevier* in its journal list compiled for the *Scopus* database, and hence imported into *SJR*.

Conceptually and linguistically there does not seem to be much of a difference between the subject terms Psychiatric Mental Health and Psychiatry and Mental Health. Still, both appear on the subject category list. The former is assigned to 18 journals, the latter to 329 journals, and 8 of them appear in both subject categories. There are even more complicated situations. For example, there are subject categories for Management Information Systems (54 journals), Information Management Systems (46 journals), Information Systems (132 journals), Health Information Management (54 journals) and Health Informatics (31 journals).

This confuses even to experienced librarians who recall from the 1970s the displeasure of searching multiple databases using phrase-indexed descriptor terms from the thesauri of half a dozen databases that assigned practically the same terms but with slightly different punctuation and structure. The search engine did not tolerate an extra space, or a semicolon instead of a colon. For example, the user who wants to know the standing of journals related to medical information systems would need to jump through many subject categories to learn about the availability of such journals because they are scattered across several subject categories for no obvious reason. With a good add-on utility these anomalies could be minimalized.

The inconsistency and sloppiness in the assignment of journals to subject categories undermines the purpose of creating disciplinary league lists of journals. Assigning journals to several categories is not a bad idea, if "there is a method to

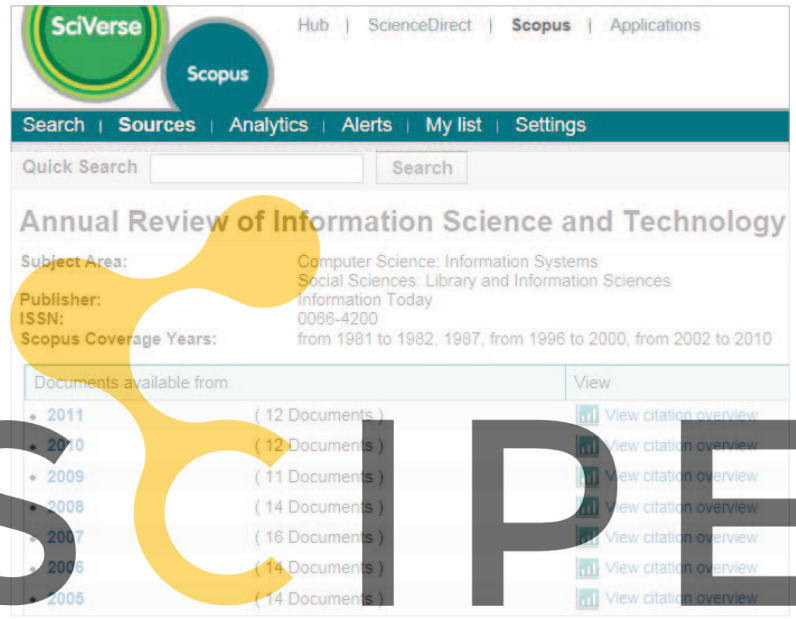


Figure 4a and 4b. Complete absence of *Arist* in *SJR*, complete coverage in *Scopus* for the last 10 years.

it". It is appropriate and informative to see that the *Health information and libraries journal* is assigned to Health Informatics, Health Information Management, and LIS, and is in the 2nd quartile in all these subject areas. Its assignment to the Medicine (Miscellaneous) subject category is questionable, and its position in the top quartile is somewhat surprising, until one realizes that there are 1,563 journals in the Medicine (Miscellaneous) subject category, and it is listed as the 360th journal by the h-index and 304th by its *SJR* score among the 390 journals in the top quartile for this subject area.

"Everything is miscellaneous"

Each subject area has a miscellaneous subject category, often with a very large number of journals, proving the adage made famous by the bestseller book that argued "everything is miscellaneous" (**Weinberger**, 2007), i.e. there is no universally agreeable method for classifying information.

The 26 subject areas and the 300+ subject categories are browsable – a very useful feature. It could be further improved by offering an option to browse the subject terms in a hierarchical way, e.g. to look up all the subject categories under the Computer Science broad subject area. The subject areas have a 4-digit code and the subject categories code have the same first 2 characters. The miscellaneous subject category is always the first one. Currently these are not used in *SJR*; they were picked up from the *Scopus* journals list for this paper. The following series (figure 3) is not screenshots from the browsable subject index but excerpts from my mock-up sheets to illustrate how the enhanced browsable subject category index should look (i.e. alphabetical and shows the category code, and the number of journals assigned to a particular subject category). Each should start with the main category code and all upper-case subject area entry, such as Computer Science. It would give the user an instant feel for the number of journals within the subject area and its subject categories. The number of journals per category should not be calculated, as a journal may be assigned to more than one of the subject categories. The average number of journals in a subject category is 102; the median number is 65 across *SJR*-2011.

The redundant subject category terms and those with no associated journals –such as Medical Terminology, Nurse Assisting- should be deleted. The 28 other categories with only a couple of journals assigned to them –such as Dental Assisting, Dental Hygiene, Periodontics, Podiatry, Optometry- should be merged into other, semantically closest, subject categories. The inadequate and unprofessional treatment of the Health Profession and Dentistry subject areas clearly stand out at the level of granularity of the earlier observation that they should not be designated among the 26 broad subject areas. They also illustrate the overuse of the miscellaneous category and the irrational subject categories with 0-3 journals (figure 3b). One would expect that *Elsevier* could have come up with a few dozen additional and more specific subject categories rather than merely parroting the name of the broad subject area with 606 journals assigned to the Economics, Econometrics and Finance subject areas (figure 3c). The thesaurus of the *EconLit* database could have given some inspiration for *Elsevier*'s indexing specialists.

The treatment of some subject categories also reinforces the subtitle of the above mentioned book ("the power of the new digital disorder"). In *Elsevier*'s journals classification, it is madness and there is **no** method to it - to paraphrase Hamlet. This is not the venue to philosophize about ontologies, but rather to recommend that *SJR* developers add a relatively simple software solution to alleviate or even solve the problems in *Elsevier*'s "ontology" for the subject classification of journals. For the *El profesional de la información* readership, the best way to help drive home these points about the inaccuracies and inconsistencies in the imported subset of the *Scopus* database may be to illustrate the implications through the prism of (re)assigning LIS journals from different subject categories in a flexible,

non-prescriptive way. The recommended software enhancements could be implemented by the *SJR* developers, and easily used by librarians and other interested end-users for any other subject categories to create more appropriate league lists for any of the subject categories.

The LIS subject category and its source base

There is a subject category in *SJR* for Library and Information Sciences. Almost all the 134 journals in it are LIS-related, although to different extent. Only three journals don't seem to be relevant for the LIS[4] subject area (which is used here in a broad sense to include Library and Information Sciences/Systems/Services/Studies). These are the *Canadian journal of program evaluation, Development and learning in organisations,* and the *International journal of hospitality and tourism administration*. Just because they may have a few papers which discuss evaluation of a LIS Program or the role of a library as a learning organization, or the comparison of the far less stellar, no tuxedo required, library facilities on cruise ships, they should not be included in the LIS league list. *Terminology* and the *International journal of lexicography* may not seem to belong to LIS, but it is appropriate to assign them to it because both journals have papers related to natural language searching, controlled vocabularies, indexing and abstracting.

It is surprising at first sight that several core LIS journals do not appear in the LIS category league list of *SJR*, such as *Journal of academic librarianship*, *Information processing & management*, *Journal of documentation*, *Online information review*, *Information technology and libraries*, *Internet research*, *Law library journal*, *Canadian journal of information and library science*, *Journal of scholarly publishing*, and *School library media research*.

The good news is that all of the above-mentioned serial sources are included in *SJR*. However, they are assigned (by *Elsevier* for the *Scopus* journal list) to the Education, Information Systems, Information Management Systems, Management Information Systems, Law, Business, Management, Accounting, Communications and Media Technology subject categories, but are not assigned (also) to the LIS category. They should be and could be, because multiple category assignment is possible, and some journals are –reasonably- assigned to more than three subject categories in the journal list of *Scopus*, and hence in *SJR*. The process of re-assigning journals to subject categories by librarians and other sufficiently interested end-users could be very easy if the *SJR* developers would add a simple utility to the *SJR* software.

The bad news is that some highly relevant and active LIS journals – especially from the School Librarianship and the Archiving fields- are not covered by *Scopus* and hence by *SJR* at all. These include *School libraries worldwide*, *Teacher-librarian*, *Journal of information literacy*, *Archivaria*, *Archives*, and *Archival issues.* Important journals from other LIS fields are also totally absent in *Scopus* and hence in *SJR*, such as *Information technologies*, *Library technology reports*, *Evidence-based library and information practice*, *New review of information behavior research*, and *Journal of education for library and information science.* Only *Scopus* can help in this regard.

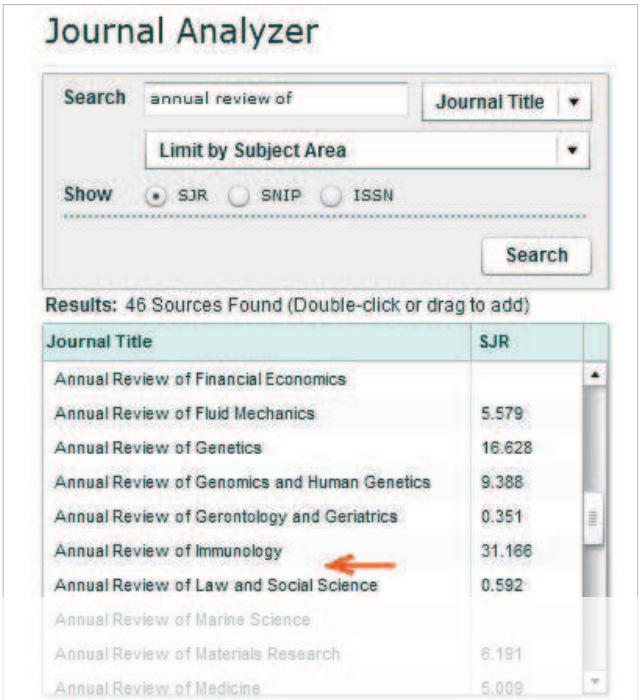


Figure 4c. Enigmatic absence of *Arist* among other annual reviews in *Scopus* Journal Analyzer

Oddly, the *Annual review of information science & technology* (*Arist*), which would have had the top ranking position by several indicators for its 2008-2011 volumes, is not included in the LIS category, nor in any other categories of the most current edition of *SJR* (figure 4a).

Scopus has had continuous coverage (at least for the past decade) of the 12-15 *Arist* chapters (figure 4b) typically issued every year, so its complete omission from the *SJR* is enigmatic. It was included in *SJR*'s earlier editions, and in a review four years ago (Jacsó, 2009b) ironically it was used as an illustration for the content-rich and very well presented *SJR* league lists. An attempt to use the Journal Analyzer module of *Scopus* to trace this oddity suggested that the problem is on the *Scopus* side, which does not show *Arist* among the 46 Annual Reviews series (figure 4c). The concern is not merely about this particular publication, but other serials which may not show up in any of the league lists for some reason.

Journals to be excluded from the *SJR*-2011 for league lists

There are more than a thousand journals that should have not been included in the 2011 edition simply because data are lacking for one or more years in the 2008-2011 volumes. That should have been a primary filtering criterion for importing *Scopus* data. Such journals do not qualify for *SJR*-2011 if the gaps were caused by inconsistent indexing in *Scopus*.

The most handicapped journals in this scenario are those for which data are not at all available for the first year of the target time period, i.e. 2008 for *SJR*-2011. Papers published in 2008 had the best chance to accumulate citations over the longest time and therefore to increase the journals' indicator scores. Practically, restricting the import to

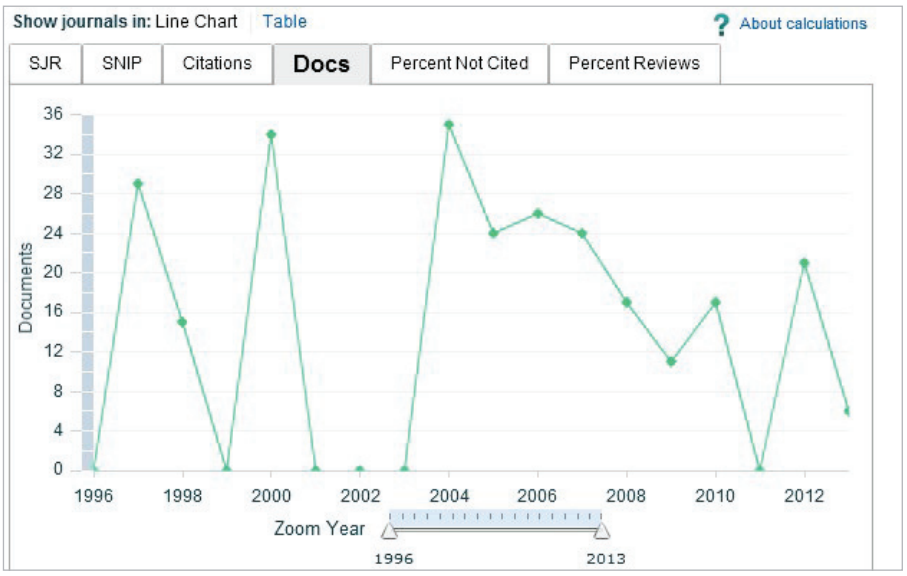


Figure 5. Multiple missing years in the run of the *International journal on digital libraries*

journals with complete data for each year in the assessed time period, i.e. 2008-2011 in case of *SJR*-2011, would be the fairest process. For the LIS league, such exclusion would reduce the number of qualifying journals by 25 and the set of LIS journals to 109.

The *SJR* indicator is based on the algorithm which calculates a size dependent score for each journal's performance. Even if "only" the data for a single year are missing for a journal, this will distort the first step of the calculation and misrepresent the indicator scores if the indexing/abstracting service (*Scopus*) missed an entire year and *SJR* inherits the incomplete data. The exception is if the journal did not publish any papers in one or more years (2001-2003, 2011) –as happened with the *International journal on digital libraries*, for example, in spite of its very appealing title and stellar advisory board. It gives the impression that the journal had a roller-coaster publishing pattern, without the fun of it (figure 5). In such case the journal should not be removed from the league list, as the problem is not caused by missing issues and volumes in *Scopus*. The technical aspects of verifying the "culprit" of such gaps will be discussed in the follow-up paper.

It may be an unpopular, "unpatriotic" suggestion from an LIS professor to remove journals from the LIS subject categories, but it could be gently done by the end-user(s) with the minor additional software features recommended below. At least 50-60 journals could be moved or added to the LIS subject category from several other categories, in addition to those that were mentioned earlier as assigned to several categories but not LIS. This is a scenario where the need to reassign the stray journals is obvious, as they undoubtedly belong to LIS, such as the *Journal of documentation*, *Online information review* or *Journal of information systems*, etc.

In this case, most of them would come from the categories of Information Systems (38), Information Systems and Management (5), and Management Information Systems (6), with a few more from Education, Communication, Media Technology, Law, Business Management and Accounting. This is an ultra-conservative count. Of the 54 journals assigned to the Management Information System category, only those with "information management" in the title, rather than "management information", were included, and of the 204 journals in the Computer Science Applications subject category only those dealing directly –in this author's experience- with information science, systems, service, studies (i.e. IS[4] analogously to LIS[4] discussed earlier) were counted.

If the time period is extended to 5 years, i.e. 2007-2011, an additional 9 journals would need to be removed from the current LIS subject category for not having any data for either or both of the two additional years, 2006 and 2007. As discussed earlier, this 5-year time span would allow a comparison with the indicators used in *JCR*-2011, *Eigenfactor*-2011, and *GSMP*-2011. Despite the removal of journals, *SJR* still would have the largest LIS set among those competitors which have the LIS category and show the entire league list of the categories, not just a teaser as *GSMP* does. Adding LIS journals scattered across many other subject categories would give *SJR* by far the largest LIS category, with about 200 active journals. The process of customizing the journal list of categories until a consensus is reached by all parties or by a supermajority of those involved in the decision-making process could be made very efficient by enhancing the *SJR* software.

Customizing the journal set of subject categories

Looking up key indicators of the status of journals is just the first step. The next one is to create a league list from the set of journals needed by the institution for collection development and for the assessment of the publishing performance (productivity and impact) of its research faculty. The option for librarians and other information professionals of modifying the journal lists in existing categories and even creating new categories, permanently or on the fly, would be very useful for producing customized league lists to match the list of preferred or expected publishing outlets of a university, college or department for decisions related to tenure, promotion and grant applications, as well as for serials collection development.

It cannot be expected that the members of the *SCImago Research Group* would fix the inconsistent and inappropriate subject classification errors found in *Scopus*, and no classification system could please all the users all the time. There still would remain journals that should be added to or removed from existing "pret-a-porter" league lists to match the "official" preferences of the different units of an institution. However, the *SCImago Research Group* could offer a built-in software solution for this problem as an add-on utility that would re-rank the league list, keeping the very appealing

layout and filtering features while providing a customized version of the *SJR* that would please most of the users most of the time at an institution that prefers to create some tailor-made disciplinary league lists from the nearly 20,000 journals.

For example, a Library Studies (LS) category created by the end-users of *SJR* could produce a shorter, much more focused list of perhaps 60-65 journals, removing IS[4] (information sciences/systems/services/studies), communications, and other journals and adding some LS journals currently in other *SJR*-2011 categories to avoid comparing apples to oranges. Similarly, creating a category for IS[4] by removing LS journals, and adding IS journals from other categories, would produce a better league list for those institutions where library science is not taught, but courses related to IS[4] are.

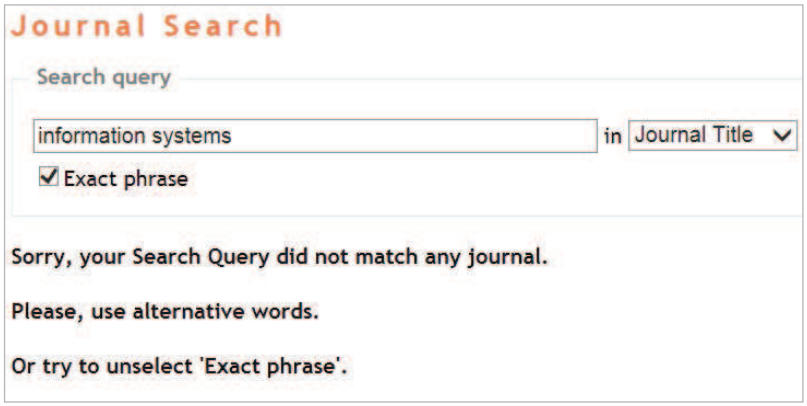


Figure 6. Dysfunctional option for exact phrase searching

The case of the *University of Hawaii* (*UH*) may well illustrate this need for upward and downward customization, i.e. broadening and narrowing the journal set. It would require somewhat different league lists for assessing the publishing performance of the faculty and/or doctoral students of the *Library and Information Sciences* (*LIS*) *Program*, the *Department of Information & Computer Sciences* (*ICS*), and the *Communication and Information Sciences* (*CIS*) interdisciplinary Ph.D. program. The latter is a joint effort of 4 units, including *LIS*, *ICS*, the *School of Communications* within the *College of Social Sciences*, and the *Department of Information Technology* in the *College of Business*. Substantial differences exist between the different colleges and departments concerning which serials are preferred.

> The *SCImago Research Group* could offer a built-in software solution for this problem as an add-on utility that would re-rank the league list

In spite of the good cooperation among the representatives of the 4 units in their regular meetings to assess the progress of doctoral students in terms of publishing efforts and achievements, the process is quite time-consuming, as expectations from the candidates differ – depending on their dissertation topic- regarding the preferred publishing venues. These meetings would be more efficient if a customized league list of the preferred journals and conference proceedings could be modified and displayed ad-hoc to check their status indicators in the context of related serials. This would be even more important at the meetings of the system-level tenure and promotion review committees (TPRC) where applications from Astrophysics to Zoology faculty are reviewed and discussed and none of the members may be sufficiently familiar with the standing of the journals in those fields. They may bring up various metrics from different services to support their pro or con arguments without knowing or explaining the limitations and biases of the widely popular, but non-transparent, services such as *GSMP*.

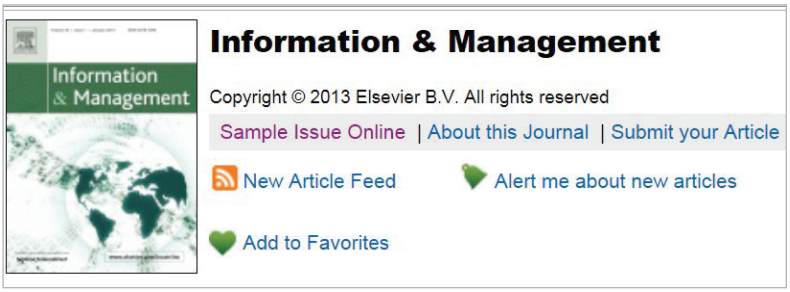
Software enhancements for customizing journal sets in *SJR*

There are two essential steps for customizing the journal sets assigned to the subject categories for the *Scopus* database and imported into *SJR*. One is to (re)move journal(s) from an inappropriate subject category, such as the *International journal of hospitality and tourism administration* from the LIS category, which needs to be moved to the Tourism, Leisure and Hospitality Management category where it is not included, so simply removing it would make the journal disappear from *SJR*. The other essential step is to locate a journal which is not included in the expected subject category, such as the *Journal of informetrics*, which is assigned to 5 categories but not to LIS, as was discussed above. Searching is critical as the users may not be expected to keep jumping from one category to several others, until they find the journal, or give up the "wild goose chase".

Locating journals in *SJR*

Finding quickly some journals that don't appear in the league list of the subject area where the user expected to find them is very important, especially considering the deficiencies in the assignment of journals to subject categories. This was the case in the test phase with the LIS category for about 60 journals that are undoubtedly related to the LIS disciplinary field, but were not listed there. The elegance and simplicity of the current user interface for the league tables, the filtering, sorting options and the league list presentation of *SJR* are as appealing as flamenco dancing by the late Antonio Gades. These should be enhanced by a search box with a smart and powerful search engine directly on the template page originally designed for looking up/browsing league tables.

There is a Journal Search module in *SJR* but it should not be used for several reasons. It is a disappointingly primitive component, as if it had been outsourced to an undergraduate student who quit the undergraduate course on Computer Programming 101 after the first session. In many regards this search module performs as poorly as elderly tourists trying their first flamenco steps while watching a show during a sightseeing tour of Seville.



Figures 7a and 7b. Metamorphosis of **&** into **and** from the cover page in *ScienceDirect* to *Scopus* and *SJR*

The inferior search module merely searches for the character string entered in the search box, bringing up odd results –actually, not searching but scanning for a character string match. For example, for the search term **tax** it retrieves *Taxon*, *Zoofaxa*, *Mycotaxon*, *Syntax*, etc., beyond the relevant journals. To find journals with the title word **reference** it also retrieves journals such as *Food quality and preferences*, *Patient preference and adherence*. There is no explicit truncation symbol, but entering the search term **libr** will find library, libraries, librarian(s), and librarianship. It can be convenient, but because of string matching rather than title word searching, it also retrieves *Fluid phase equilibria*, *Equilibrium research*, and several other journals. The term **online** finds a number of journals with nonlinear in their title, such as *Nonlinear dynamics, Nonlinearity, Journal of nonlinear science*. Users would still feel that they went for a wild-goose chase as *Online*, the excellent LIS trade journal, is not available in *SJR* even if it has been covered well in *Scopus*.

There is an exact phrase search option using a check box, but it does not work. With the check box activated, the search term **information systems** should find only the journal with that two-word title without additional word(s), but whenever the check box is activated it triggers an error message. Unselecting the exact phrase option will produce a result list of 32 journals which match the character string, but include additional word(s). Searching by ISSN will bring back the wanted journal, but it is not an efficient solution when trying to locate quickly the journals with the term **information systems** in their title, as the result list is sorted by the h-index and cannot be re-sorted by title. Searching by publisher will work, but the data is not current (reflecting the time before *Elsevier*, *Taylor and Francis*, *Springer* went on a shopping spree and acquired dozens of publishers), and for league list purposes this metadata is of no primary importance.

There are no Boolean operators. It is impossible to formulate queries like **age OR aging OR aged** and it adds insult to injury that the string-matching module retrieves many journals with the words image, imaging, language, management, manager, managing, managed, etc. in their title. The query **information AND management** will retrieve two journals where the title includes exactly that 3-lettter character string, such as *Information and management* (although the journal itself uses the "&" sign). The journal *Information processing & management* is not retrieved because information and management are not adjacent words in the journal's title. In other words, AND is not a Boolean AND operator but a character string which should appear in upper or lower case in the journal title to match it. The search term **Information & Management** does not yield any result because the '&' symbol is transformed to the character string 'and' in *Scopus*, and *SJR* inherits this spelling "feature" (figures 7a and 7b). These bad traits of the search module –which hinder the *SJR* software- are reminiscent of the first efforts of library automation in the former Soviet Union in the early 1960s.

No sorting option is offered for results produced by the current search module. The results lists are by default sorted by the h-index of the journals. This is not clear for the typical end-user because the h-index is not displayed. A good search module should offer all of the sort options available in the league list module (Title, *SJR* score, Total documents, Total cites, Citable documents and Cites/Documents). A smart search engine box located on the league table template could replace the inferior search module. The only feasible way to implement the recommended changes is to empower even the casual user to customize the list and feel/look like a regular ballroom dancer.

Filtering and sorting the search results

The *SJR* league list template offers good options to restrict the search beyond the subject categories. The option for filtering by the country of the journal publishers' headquarters is not really as relevant as in the *SCR* (*SCImago Country Report*) database, where country refers to the authors' affiliation(s), and can provide an approximation (in the context of other countries) for the productivity and impact of the document published by authors' affiliation. In *SJR* the country name refers to the journal's place of publication.

Limiting by language of publication could be useful because end-users may want to limit the search in this way (e.g., to only English and/or Spanish language journals), to exclude journals in the other 50+ languages which are unknown and irrelevant for the vast majority of the researchers, universities and libraries around the world. The other option for finding journals is by ISSN and publisher name. Both of these options should be kept in the new smart search engine on the league list template, rather than as a separate function as it is now.

Adding, moving and removing journals as an end-user function

The results list of the league tables should be enhanced in the new search engine by offering an ADD TO, MOVE and REMOVE button next to each item, allowing end-users to make a customized league list to their liking. End-users could just click on the appropriate button and choose from a scroll-down list the subject categories to add or remove the journal, such as adding the *Journal of documentation* to the LIS category, or removing *Development and learning in organisations* from the LIS category. Removing journals from their current category would just require a click on the check-box. It would be the responsibility of the end-users to realize when removing (i.e. deleting) is not the appropriate option for a journal that is assigned only to a single subject category, but a wrong one, as in the case of the *Canadian journal of program evaluation*, which appears only under the LIS category. In such case it must be moved to the appropriate category. It is important to have an optional log-in function or a cookie to keep the customized league lists for later use/updating.

The competent and careful re-assignment of journals to subject categories can make the league lists more fair and credible from the perspective of editors and authors. There is nothing wrong with assigning a journal to more than one subject category, but mis-assignment of journals based solely on *Elsevier*'s judgment can unjustly handicap a journal, because its rank position depends much on the company of the journals in the debatable category, especially if that is the only category to which the journal is assigned. The journal keeps the same indicator set independent of the category where it ends up, but it may be similar to assigning a boxer to the Bantamweight category when he qualifies to the Flyweight category.

A summary was created to compare the key metrics indicators of the current LIS category and of the selected subset of 38 journals from the 182 journals assigned to the Information Systems (ISY) category and intended to be added to the LIS category. The mean, median and maximum scores clearly indicate that the journals in the ISY category have significantly higher scores, more like Flyweight versus Heavyweight contenders. If an LIS journal is assigned to the ISY category only, its rank position is much handicapped by all eight metrics.

Adding the 38 journals from the ISY category to the LIS category would produce better rank positions for them. For example, the *Journal of electronic publishing* currently assigned to the ISY subject category within the Computer Science broad subject area is in the bottom quartile (Q4) by *SJR* score; it would be in the lower quartile (Q3) if it were moved or added to the LIS category within the Social Sciences broad subject area. The same is true for the *Canadian journal of information and library science*. The one quartile change is from Q3 to Q2 (upper quartile) for *Technical services quarterly.*

SJR uses Q1 for top, Q2 for upper, Q3 for lower and Q4 for bottom quartile, which is a less common notation than the reverse (Q4 for top, Q3 for upper quartile, etc.). A flip-flop option or a cookie setting notation for this as well as for the decimal fraction and thousand separator notation from the Continental to the US system would be useful as part of the one-time customization process.

For most of the indicators, a similar option to change from the quartile to the percentile system could provide a closer distinction among the journals' standings within a league list as another layout option for those who are less interested in the absolute unit values than in the easily interpreted, directly comparable percentile ranks. Showing the measures of central tendency and the range on the top of the league list could also offer a set of traditional statistical indicators that would help the user get a better feel for the context.

Conclusions

SJR is the largest open access directory, covering nearly 20,000 journals and other serial publications. It is among the best-designed services, with an intuitive interface, browsable indexes, sorting and filtering options, and compact but information-rich journal league lists in 300 subject areas. The entire content can be downloaded in one fell swoop into an Excel file for further processing by any spreadsheet software.

It needs a far better search engine to help in finding journals that are not always assigned by *Elsevier* to the most reasonable, most likely or expected categories. A journal may be ranked far lower by many of the metrics indicators in one subject category than in the other much more appropriate category because of the syndrome of comparing apples to oranges. End-users should be empowered to customize the

Figure 8. Measures of central tendency for key metrics for the journal in the LIS versus the ISY subject categories

Title	SJR score	H index	Total docs. (2011)	Total docs. (3 yrs)	Total refs. 2011	Total cites (3 yrs)	Cites/Doc. (2 yrs)	Refs./Doc
Library and Information Science (LIS) alone: 134 journals								
mean	0.419	11	37	105	891	89	0.65	24.13
median	0.276	6	29	82	523	30	0.44	22.74
max	3.652	78	274	805	9,370	1,614	4.01	61.91
Information Systems alone: 38 journals								
mean	1.058	20	50	123	2,111	349	2.04	46.74
median	0.555	14	29	69	1,045	95	1.50	38.92
max	10.411	94	409	1,051	14,892	4,670	14.43	263.00

journal lists easily online to meet their expectations/preferences, while retaining all the existing browsing and search features and the superb design and implementation of the *SJR* layout. It has given the Spanish *SCImago Research Development Group*, and LIS community, a good reputation. Relatively minor software enhancements by the software developers could make the service even more popular in collection development and in metrics-based research evaluation through proxy movement, where the league list position of the publishing outlets of individuals, groups, departments and entire universities strongly influence decisions related to career advancements, accreditation and funding applications.

The ontology of *SJR* subject areas and categories has weak points and shows negligence, as does the assignment of journals to subject categories. These are inherited from the content and structure of the journal list compiled by *Elsevier* for the *Scopus* database, which has had many innovative software features and many shortcomings in content. There are still significant gaps in the breadth of coverage of the source publications for *Scopus*, which are inherited by *SJR*. These may have a significant effect on the rank position of journals. A follow-up paper will discuss this issue.

References

Bar-Ilan, Judit (2010). "Rankings of information and library science journals by JIF and by h-type indices". *Journal of informetrics*, v. 4, n. 2, pp. 141-147.
http://dx.doi.org/10.1016/j.joi.2009.11.006

Bergman, Elaine M. Lasda (2012). "Finding citations to social work literature: the relative benefits of using *Web of Science, Scopus*, or *Google Scholar*". *The journal of academic librarianship*, v. 38, n. 6, pp. 370-379.
http://dx.doi.org/10.1016/j.acalib.2012.08.002

De-Moya-Anegón, Félix; **Chinchilla-Rodríguez, Zaida**; **Vargas-Quesada, Benjamín**; **Corera-Álvarez, Elena**; **Muñoz-Fernández, Francisco-José**; **González-Molina, Antonio**; **Herrero-Solana, Víctor** (2007). "Coverage analysis of Scopus: A journal metric approach". *Scientometrics*, v. 73, n. 1, pp. 53-78.
http://dx.doi.org/10.1007/s11192-007-1681-4

Deis, Louise; **Goodman, David** (2005). "*Web of Science* (2004 version) and *Scopus*". *The Charleston advisor*, v. 6, n. 3, pp. 5-21.

Delgado-López-Cózar, Emilio; **Cabezas-Clavijo, Álvaro** (2012). "*Google Scholar Metrics*: an unreliable tool for assessing scientific journals". *El profesional de la informacion*, v. 21, n. 4, pp. 419-427.
http://dx.doi.org/10.3145/epi.2012.jul.15

Delgado-López-Cózar, Emilio; **Cabezas-Clavijo, Álvaro** (2013). "Ranking journals: could *Google Scholar Metrics* be an alternative to *Journal Citation Reports* and *SCImago Journal Rank*?". *Learned publishing*, v. 26, n. 2, pp. 101-113.
http://dx.doi.org/10.1087/20130206

Dess, Howard M. (2006). "Database reviews and reports. Scopus". *Issues in science and technology librarianship*, Winter.
http://www.istl.org/06-winter/databases4.html

Falagas, Matthew E.; **Kouranos, Vasilios D.**; **Arencibia-Jorge, Ricardo**; **Karageorgopoulos, Drosos E.** (2008). "Comparison of SCImago Journal Rank indicator with Journal Impact Factor". *The Faseb journal*, v. 22, n. 8, pp. 2623-2628.
http://dx.doi.org/10.1096/fj.08-107938

Garfield, Eugene (1999). "Journal impact factor: a brief review". *Canadian Medical Association Journal*, v. 161, n. 8, pp. 979-980.
http://www.garfield.library.upenn.edu/papers/journalimpactCMAJ1999.pdf

Garfield, Eugene (2005). "The agony and the ecstasy - The history and the meaning of the Journal Impact Factor". In: *The intl congress on peer review and biomedical publication*. Chicago, USA. September 16.
http://www.garfield.library.upenn.edu/papers/jifchicago2005.pdf

Gómez-Núñez, Antonio J.; **Vargas-Quesada, Benjamín**; **De-Moya-Anegón, Félix**; **Glänzel, Wolfgang** (2011). "Improving SCImago Journal & Country Rank (SJR) subject classification through reference analysis". *Scientometrics*, v. 89, n. 3, pp. 741-758.
http://dx.doi.org/10.1007/s11192-011-0485-8

González-Pereira, Borja; **Guerrero-Bote, Vicente P.**; **De-Moya-Anegón, Félix** (2009). "The SJR indicator: A new indicator of journals' scientific prestige".
http://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf

González-Pereira, Borja; **Guerrero-Bote, Vicente P.**; **Moya-Anegón, Félix** (2010). "A new approach to the metric of journals' scientific prestige: The SJR indicator". *Journal of informetrics*, v. 4, n. 3, pp. 379-391.
http://dx.doi.org/10.1016/j.joi.2010.03.002

González-Pereira, Borja (2011). Personal communication.

Jacsó, Péter (2007). "Scopus", *Gale: Péter's digital reference shelf*, 2007, November.
http://jacso.info/PDFs/Jacso-Scopus-2007-11-Gale.pdf

Jacsó, Péter (2008). "The pros and cons of computing the h-index using Scopus". *Online information review*, v. 32, n. 4, pp. 524-535.
http://dx.doi.org/10.1108/14684520810897403

Jacsó, Péter (2009a). "Errors of omission and their implication for computing scientometric measures in evaluating the publishing productivity and impact of countries". *Online information review*, v. 33, n. 2, pp. 376-385.
http://dx.doi.org/10.1108/14684520910951276

Jacsó, Péter (2009b). "SCImago, Worldmapper, Atlapedia". *Online*, v. 33, n. 3, pp. 50-54.

Jacsó, Péter (2009c). "Scopus", *Gale: Péter's digital reference shelf*, June.
http://www.gale.cengage.com/reference/peter/200906/scopus.html

Jacsó, Péter (2010). "Comparison of journal impact rankings in the SCImago Journal & Country Rank and Journal Citation Reports databases". *Online information review*, v. 34, n. 4, pp. 642-657.
http://dx.doi.org/10.1108/14684521011073034

Jacsó, Péter (2012). "Google Scholar Metrics for Publications – The software and content feature of a new open access bibliometric service". *Online information review*, v. 36, n. 4, pp. 604-609.
http://dx.doi.org/10.1108/14684521211254121

Jacsó, Péter (2013). "The role of information professionals in interpreting the results of journal league lists for research performance evaluation" (Prologue), *Anuario ThinkEPI*, v. 7, pp. 11-15.
http://thinkepi.net

Moed, Henk F. (2010). "Measuring contextual citation impact of scientific journals". *Journal of informetrics*, v. 4, n. 3, pp. 265-277.
http://dx.doi.org/10.1016/j.joi.2010.01.002

Oppenheim, Charles (2008). "Out with the old and in with the new: the RAE, bibliometrics and the new REF". *Journal of librarianship and information science*, v. 40, n. 3, pp. 147-149.
http://dx.doi.org/10.1177/0961000608092550

Pudovkin, Alexander I.; **Garfield, Eugene** (2004). "Rank-normalized impact factor: A way to compare journal performance across subject categories". *Proceedings of the ASIST annual meeting*, v. 41, pp. 507-515.

Pudovkin, Alexander I.; **Garfield, Eugene** (2009). "Percentile rank and author superiority indexes for evaluating individual journal articles and author's overall citation performance". Presented at: *Fifth intl conf on webometrics & scientometrics (WIS)*. Tenth Collnet Meeting, Sept.13-16, Dalion, China.

Raj, Ram-Gopal; Zainab, Awang N. (2012). "Relative measure index: a metric to measure the quality of journals". *Scientometrics*, v. 93, n. 2, pp. 305-317.
http://dx.doi.org/10.1007/s11192-012-0675-z

Torres-Salinas, Daniel; **Jiménez-Contreras, Evaristo** (2010). "Introducción y estudio comparativo de los nuevos indicadores de citación sobre revistas cientificas en *Journal Citation Reports* y *Scopus*". *El profesional de la información*, v. 19, n. 2, pp. 201-207.
http://dx.doi.org/10.3145/epi.2010.mar.12

Wagner, A. Ben (2009). "Percentile-based journal impact factors: a neglected collection development metric". UB Faculty and Staff Publications.
http://ubir.buffalo.edu/xmlui/handle/10477/1318

Waltman, Ludo; **Van-Eck, Nees-Jan**; **Van-Leeuwen, Thed N.**; **Visser, Martijn S.** (2012). "Some modifications to the SNIP journal impact indicator". *Journal of informetrics*, v. 7, n. 2, pp. 272-285.
http://dx.doi.org/10.1016/j.joi.2012.11.011
http://arxiv.org/ftp/arxiv/papers/1209/1209.0785.pdf

Weinberger, David (2007). *Everything is miscellaneous: the power of the new digital disorder.* Henry Holt and Company, LLC. ISBN: 0 8050 8811 3